

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**



(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平6-222797

(43)公開日 平成 6年(1994) 8月12日

(51)Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/14	G	8946-5H		
	J	8946-5H		
9/18	E	8946-5H		

審査請求 有 請求項の数 5 O L (全 15 頁)

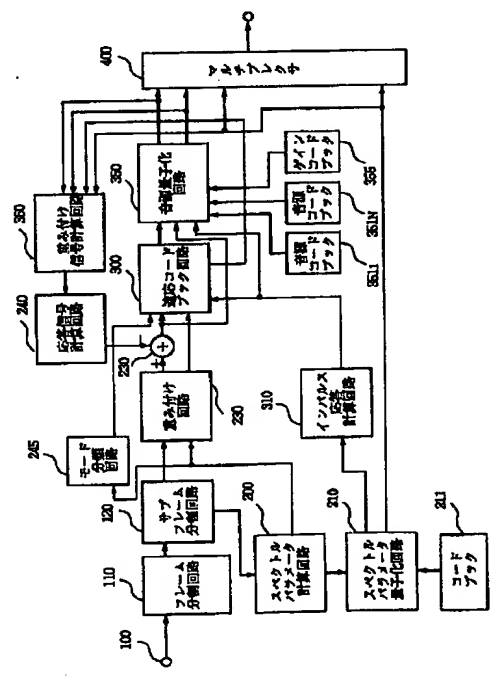
(21)出願番号	特願平5-8737	(71)出願人	000004237 日本電気株式会社 東京都港区芝五丁目7番1号
(22)出願日	平成 5年(1993) 1月22日	(72)発明者	小澤 一範 東京都港区芝五丁目7番1号日本電気株式 会社内
		(74)代理人	弁理士 京本 直樹 (外 2名)

(54)【発明の名称】 音声符号化方式

(57)【要約】 (修正有)

【目的】 4. 8 k b / s 以下のビットレートで高い音質を得る音声符号化方式。

【構成】 音声信号をフレームに分割し、サブフレームに分割し、サブフレームのスペクトルパラメータ計算部200と、スペクトルパラメータを量子化するスペクトルパラメータ量子化部210と、音声信号から特徴量を計算してフレーム音声信号のモード分類部245と、スペクトルパラメータを用い音声信号に聴感重み付けを施した重み付け部230と、モード分類結果と量子化されたスペクトルパラメータと音声信号とを用いて、ピッチ周期を表わすパラメータを求める適応コードブック部300と、重み付け信号と適応コードブック部300の出力とスペクトルパラメータと量子化されたスペクトルパラメータを用い、複数段からなる音源コードブック351₁~351_Nとゲインコードブック355とを探索し音源信号を量子化する音源量子化部350とを有する。



【特許請求の範囲】

【請求項1】 入力する音声信号をあらかじめ定めたタイミングごとにフレームに分割しさらに複数個のサブフレームに細分割して、少なくとも一つの前記サブフレームに対して前記音声信号のスペクトルの特徴を表すスペクトルパラメータを算出するスペクトルパラメータ計算部と、あらかじめ指定した位置の前記サブフレームのスペクトルパラメータを量子化コードブックを用いて量子化するスペクトルパラメータ量子化部と、前記音声信号のあらかじめ定められた特徴量を計算して前記フレームの音声信号を複数種類のモードのうちの一つに分類するモード分類部と、前記スペクトルパラメータ計算部にて得られる前記スペクトルパラメータに応じて前記音声信号に聴感重み付けを施した重み付け信号を求める重み付け部と、前記モード分類部でのモード分類結果と前記スペクトルパラメータ計算部にて得られるスペクトルパラメータと前記スペクトルパラメータ量子化部にて量子化されたスペクトルパラメータと前記重み付け信号とに応じて、そのモードに対応した前記音声信号のピッチを表わすパラメータを求めて送出する適応コードブック部と、前記重み付け信号と前記適応コードブック部の送出パラメータと前記スペクトルパラメータと前記量子化されたスペクトルパラメータとに応じて、複数段からなる音源コードブックとゲインコードブックとを探索して前記音声信号の音源信号を量子化した出力信号を発する音源量子化部とを有することを特徴とする音声符号化方式。

【請求項2】 前記モード分類部は、前記重み付け部にて得る前記重み付け信号からサブフレーム毎にピッチ予測による歪を求め、そのピッチ予測歪をフレーム全体にわたり累積した値を特徴量として用いてモード分類を行う請求項1記載の音声符号化方式。

【請求項3】 前記スペクトルパラメータ量子化部は、前記モード分類部におけるモード分類結果に応じて複数段の前記量子化コードブックを切り替えてスペクトルパラメータを量子化する請求項1記載の音声符号化方式。

【請求項4】 前記音源量子化部は、前記モード分類部におけるモード分類結果に応じて複数段の前記音源コードブックと前記ゲインコードブックとを切り替えて前記音源信号を量子化する請求項1記載の音声符号化方式。

【請求項5】 前記音源量子化部は、複数段の前記音源コードブックのうち、少なくとも一つの段があらかじめ定められた間引き率のコードブックから構成されている請求項1記載の音声符号化方式。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、音声信号を低いビットレート、特に4.8kb/s以下のビットレートで高品質に符号化するための音声符号化方式に関する。

【0002】

【従来の技術】 従来、音声信号を4.8kb/s以下の低いビットレートで符号化する方式としては、例えばシュレーダ(M. Schroeder)およびアタル(B. Atal)による“コードーエキサイテッド・リニア・プレディクション：ハイ・クオリティ・スピーチ・アト・ベリ・ロウ・ビット・レイツ(Code-excited linear prediction: High quality speech at very low bit rates)”、音響、音声および信号処理に関する国際会議の議事録(Proc. ICASSP), 1985年937ないし940ページの論文(文献1)や、クレイジン(Kleijn)らによる“インブルーブド・スピーチ・クオリティ・アンド・エフィシエント・ベクトル・クォンタイゼーション・イン・エスイーエルビー(Improved speech quality and efficient vector quantization in SELP)”、音響、音声および信号処理に関する国際会議の議事録(Proc. ICASSP), 1988年、155ないし158ページの論文(文献2)に記載されているCELP(Code Excited LPC Coding)方式が知られている。この方式では、送信側にてフレーム毎(例えば20ms)の音声信号を線形予測(LPC)分析して音声信号のスペクトル特性を表すスペクトルパラメータを抽出し、フレームをさらにサブフレーム(例えば5ms)に分割してサブフレーム毎に過去の音源信号をもとに適応コードブックにおけるパラメータ(遅延パラメータおよびゲインパラメータ)を抽出し、適応コードブックによりサブフレームの音声信号をピッチ予測し、ピッチ予測して求めた残差信号に対して、予め定められた種類の雑音信号からなる音源コードブック(ベクトル量子化コードブック)から最適音源コードベクトルを選択し最適なゲインを計算する。最適音源コードベクトルの選択は、選択した雑音信号により合成した信号と、前述の残差信号との誤差電力を最小化するように行う。そして選択された音源コードベクトルの種類を表すインデックスおよび最適ゲインならびに、前述のスペクトルパラメータおよび適応コードブックから抽出したパラメータを伝送する。受信側の説明は省略する。

【0003】

【発明が解決しようとする課題】 上述した文献1および2の従来方式では、良好な音質を得るためには音源コードブックのサイズが十分大きい(例えば10ビット)必要があった。このため、音源コードブックの探索に膨大な演算量を必要とした。さらに必要とするメモリ容量も膨大であり(例えば10ビット40次元の場合には、40Kワードのメモリ容量)、ハードウェアをコンパクトに実現するのが困難であった。また、ビットレートを低減するためにフレーム長、サブフレーム長を増大し、音

源コードブックのビット数を低減せずに次元数を増大すると、演算量はきわめて顕著に増加するという問題点がある。

【0004】コードブックのサイズを低減する方法として、例えばジュアング(B. Juang)らによる“マルチプル・ステージ・ベクトル・クオンタイゼーション・フォア・スピーチ・コーディング(Multiple Stage vector quantization for speech coding)”、音響、音声および信号処理に関する国際会議の議事録(Proc. ICASSP), 1982年、597ないし600ページの論文(文献3)に記載されているように、コードブックを多段に分割して構成し、各コードブックを独立に探索する多段ベクトル量子化法が知られている。この方法では、コードブックが複数段に分割されているので、1段あたりのコードブックのサイズは、例えばB/Lビット(ここでBは全体のビット数、Lは段数)に低減されるため、コードブック探索に要する演算量は、L段全体でも、Bビット1段に比べ、 $L \times 2^{B/L}$ に低減される。また、コードブック格納に必要なメモリ容量も同様に低減される。しかし、この方法では、各段のコードブックを独立に学習、探索しているので、Bビット1段に比べ、性能は大きく低下するという問題点がある。

【0005】本発明の目的は、上述した問題点を解決し、比較的少ない演算量及びメモリ容量により、低ビットレート特に4.8kb/s以下のビットレートで音質の良好な音声符号化方式を提供することにある。

【0006】

【課題を解決するための手段】本発明の方式は、入力する音声信号をあらかじめ定められたタイミングごとにフレームに分割しさらに複数個のサブフレームに細分割して、少なくとも一つの前記サブフレームに対して前記音声信号のスペクトル特徴を表すスペクトルパラメータを算出するスペクトルパラメータ計算部と、あらかじめ指定した位置の前記サブフレームのスペクトルパラメータを量子化コードブックを用いて量子化するスペクトルパラメータ量子化部と、前記音声信号のあらかじめ定められた特徴量を計算して前記フレームの音声信号を複数種類のモードのうちの一つに分類するモード分類部と、前記スペクトルパラメータ計算部にて得られる前記スペクトル

ルパラメータに応じて前記音声信号に聴感重み付けを施した重み付け信号を求める重み付け部と、前記モード分類部でのモード分類結果と前記スペクトルパラメータ計算部にて得られるスペクトルパラメータと前記スペクトルパラメータ量子化部にて量子化されたスペクトルパラメータと前記重み付け信号とに応じて、そのモードに対応した前記音声信号のピッチを表わすパラメータを求めて送出する適応コードブック部と、前記重み付け信号と前記適応コードブック部の送出パラメータと前記スペクトルパラメータと前記量子化されたスペクトルパラメータとに応じて、複数段からなる音源コードブックとゲインコードブックとを探索して前記音声信号の音源信号を量子化した出力信号を発生する音源量子化部とを有する。

【0007】

【作用】本発明による音声符号化方式の作用を示す。

【0008】音声信号をフレーム(例えば40ms)に分割し、さらにサブフレーム(例えば8ms)に分割する。スペクトルパラメータ計算部では、少なくとも一つのサブフレーム(例えば5つのサブフレームのうち第1, 第3, 第5のサブフレーム)に対して、周知のLPC分析を施してスペクトルパラメータ(LPCパラメータ)を求める。スペクトルパラメータ量子化部では、あらかじめ定められたサブフレーム(例えば第5サブフレーム)に対応するLPCパラメータを量子化コードブックを用いて量子化する。ここで、コードブックとしては、ベクトル量子化コードブック、スカラー量子化コードブック、ベクトル・スカラー量子化コードブックのいずれをも用いることができる。

【0009】次に、フレームの音声信号からあらかじめ定められた特徴量を計算し、この値とあらかじめ定められたしきい値とを比較して、フレーム毎に複数種類(例えば4種類)のモードに分類する。次に、聴感重み付け部では、第1, 第2, 第5のサブフレームのスペクトルパラメータ a_i ($i=1$ ないし P)を用いて、下式

(1)により聴感重み付け信号をサブフレーム毎に計算する。ただし、第2及び第4サブフレームのスペクトルパラメータについてはそれぞれ、例えば第1, 第3のサブフレーム、及び第3第5のサブフレームのスペクトルパラメータを直線補間することにより得られる。

【0010】

$$X_w(z) = x(z) \cdot \left[1 - \sum_{i=1}^P a_i z^{-i} \right] / \left[1 - \sum_{i=1}^P a_i \gamma^i z^{-i} \right] \quad \dots(1)$$

【0011】ここで、 $x(z)$ 、 $X_w(z)$ は、それぞれ、フレームの音声信号及び聴感重み付け信号の z 変換である。 P はスペクトルパラメータの次数である。また、 γ は聴感重み付け量を制御するための定数であり、通常0.8程度に選ぶ。

【0012】次に、適応コードブック部では、聴感重み付け信号に対して、サブフレーム毎に、ピッチに関する

パラメータとして遅延 T とゲイン β を計算する。ここで遅延はピッチ周期に対応する。適応コードブックのパラメータの計算法は、前記文献2を参照できる。また、特に女性話者に対する適応コードブックの性能を改善するために、サブフレーム毎の遅延をサンプリング時刻毎の整数値ではなくて小数値で表すこともできる。具体的には、クルーン(P. Kroon)およびアタル(B. A

tal) による "Pitch・プレディクターズ・ウィズ・ハイ・テンポラル・レゾリューション (Pitch predictors with high temporal resolution)"、音響、音声および信号処理に関する国際会議の議事録 (Proc. ICASSP) 1990年、661ないし664ページの論文 (文献4) 等を参照することができる。このようにすることにより、例えばサブフレーム毎の遅延量を整数値で表すと7ビット必要であったのが、小数値にすることにより8ビット程度に増加するが、女性音に関しては著し*

*く音質が改善される。

【0013】さらに、適応コードブックのパラメータの計算に関する演算量を低減化するために、まず、聴感重み付け信号に対して、サブフレーム毎に遅延の候補をオープンループ探索により、下式(2)を最大化するものから順に複数種類求める。

$$D(T) = P^2(T) / Q(T) \quad (2)$$

但し、

$$【0015】$$

$$P(T) = \sum_{n=0}^{N-1} x_w(n) x_w(n-T) \quad \dots(3)$$

$$Q(T) = \sum_{n=0}^{N-1} x_w(n-T)^2 \quad \dots(4)$$

【0016】である。以上により、サブフレーム毎に、オープンループ探索により、少なくとも1種類の遅延候補を求め、その後、過去のフレームの駆動音源信号を用いたクローズドループ探索により、サブフレーム毎に、前記候補の近傍を探索して、ピッチ周期(遅延)とゲインを求める。(具体的な方法に関しては、例えば特願平3-103262号(文献5)等を参照)

適応コードブックの遅延量は、有声区間では、サブフレーム間で極めて相関が高いため、サブフレーム間で遅延量の差分をとり、差分を伝送することにより、適応コードブックの遅延を伝送するために必要な伝送情報量を、遅延量をサブフレーム毎に独立に伝送する方法に比べ大幅に低減化することができる。例えば、フレーム毎に第1サブフレームでは、遅延量を小数値で表して8ビット※

※で伝送し、第2～5サブフレームでは、直前サブフレームとの遅延量の差分を例えば3ビットで伝送すれば、全てのサブフレームで8ビットを伝送する場合に比べ、伝送情報量をフレーム当たり40ビットから20ビットに低減化することができる。

【0017】次に、音源量子化部では、複数段のベクトル量子化用のコードブックから構成される音源コードブックを探索し、前述の重み付け信号と音源コードブック中の各コードベクトルにより重み付け再生された信号との誤差電力を最小化するように、コードベクトルが各段毎に選択される。例えば音源コードブックが2段のコードブックから構成されていると、コードベクトルの探索は下式(5)に従って行う。

$$【0018】$$

$$D = \sum_{n=0}^{N-1} [x_w(n) - \beta v(n-T) * h_w(n) - \gamma_1 c_{1j}(n) * h_w(n) - \gamma_2 c_{2j}(n) * h_w(n)]^2 \quad \dots(5)$$

【0019】上式において、 $\beta v(n-T)$ は適応コードブック部のクローズドループ探索において計算された適応コードベクトルであり、 β は適応コードベクトルのゲインである。 $c_{1j}(n)$ 、 $c_{2j}(n)$ はそれぞれ、1段目、2段目のコードブックのj番目のコードベクトル

★ル、i番目のコードベクトルを表す。また、 $h_w(n)$ は下式(6)の重み付けフィルタの特性を表すインパルス応答である。また、 γ_1 、 γ_2 はそれぞれ、1段目、2段目のコードブックに関する最適ゲインである。

$$★40 \quad 【0020】$$

$$H_w(z) = [1 - \sum_{i=1}^p a_i z^{-i} / 1 - \sum_{i=1}^p a_i \gamma_i' z^{-i}] [1 / 1 - \sum_{i=1}^p a_i' \gamma_i' z^{-i}] \quad \dots(6)$$

【0021】ここで、 γ は式(1)の聴感重み付けを制御する定数である。

【0022】次に、音源コードブックの式(5)を最小☆

☆化するコードベクトルを探索した後に、下式(7)を最小化するようにゲインコードブックを探索する。

$$【0023】$$

$$D = \sum_{n=0}^{N-1} [x_w(n) - \beta v(n-T) * h_w(n) - \gamma_1 c_{1j}(n) * h_w(n) - \gamma_2 c_{2j}(n) * h_w(n)]^2 \quad \dots(7)$$

【0024】ここで γ 、 k 、 γ_{ik} は、2次元のゲインコードブックの k 番目のゲインコードベクトルを示す。

【0025】音源コードブックの最適コードベクトルを探索するときの演算量を低減化するために、各段毎に、複数種類の音源コードベクトルの候補（例えば、1段目は m_1 種類、2段目は m_2 種類）を選択し、選択後に、1段目と2段目の候補の全組み合わせ（ $m_1 \times m_2$ ）を探索して、式（5）を最小にする候補の組み合わせを選択してもよい。

【0026】また、ゲインコードブックを探索するとき10は、上記音源コードベクトルの候補の全ての組み合わせ、あるいは、組み合わせのうち、誤差電力の小さい順に、あらかじめ定められた個数だけ選択した候補の組み*

$$D_1 = \sum_{n=0}^{N-1} x_{w1}^2(n) - P_1^2(T) / Q_1(T) \quad \dots(8)$$

【0029】ここで1はサブフレーム番号である。そして下式（9）によりフレーム全体の累積予測誤差電力を求め、この値とあらかじめ定められたしきい値とを比較※

$$D = 1 / M \sum_{i=1}^M D_i \quad \dots(9)$$

【0031】例えば、モードを4種類設けるとすると、しきい値を3種類設け、式（9）の値を3種類のしきい値と比較してモード分類を行う。なお、ピッチ予測歪としては、上記以外にピッチ予測ゲイン等を用いることもできる。

【0032】請求項3に記載した本発明の実施態様のスペクトルパラメータ量子化部では、モード分類部で分類されたモードのいくつかに対して、トレーニング信号に対して、あらかじめスペクトル量子化コードブックを作成しておき、符号化するときはモード情報を用いてスペクトル量子化コードブックを切り替えて使用する。このようにするとコードブックを格納するメモリ容量は、切り替える種類だけ増大するが、全体合計ではより大きなサイズのコードブックを有することと等価となり、伝送情報量を増大させずに性能を向上させることができる。

【0033】請求項4に記載した本発明の実施態様の音源量子化部では、あらかじめ定められたモード毎で異なる音源コードブック、及びゲインコードブックを作成しておき、符号化するときは、モード情報を用いて、音源コードブック及びゲインコードブックを切り替えて使用する。このようにするとコードブックを格納するメモリ容量は、切り替える種類だけ増大するが、全体合計ではより大きなサイズのコードブックを有することと等価となり、伝送情報量を増大させずに性能を向上させることができる。

【0034】さらに請求項5に記載した本発明の実施態様の音源量子化部では、複数段のコードブックのうちの

* 合わせに対して、式（7）に従いゲインコードブックを探索し、誤差電力を最小化するようなゲインコードベクトルと音源コードベクトルの組み合わせを求めるようにしてもよい。このようにすると演算量は増大するが、性能は改善される。

【0027】次に、本発明の特許請求の範囲の請求項2に記載した実施態様のモード分類部では、特徴量として累積ピッチ予測歪を用いる。まず、適応コードブック部のオープンループ探索によりサブフレーム毎に選択されたピッチ周期候補 T に対して、ピッチ予測歪として、下式（8）に従いピッチ予測誤差歪をサブフレーム毎に求める。

【0028】

※して、複数種類のモードに分類する。

【0030】

少なくとも一つの段は、コードベクトルの要素があらかじめ定められた間引き率（例えば、間引き率=2）のレギュラーパルス構成となっている。ここで、間引き率=1とすれば通常の構成になる。このような構成にすることにより、音源コードブックの格納に必要なメモリ量を1/間引き率、（例えば間引き率=2とすると1/2）に低減できる。また、音源コードブック探索に必要な演算量もほぼ、1/間引き率以下に低減化できる。さらに、音源コードベクトルの要素を間引いてパルス化することにより、特に音声の母音部等では、聴感上重要なピッチパルスをより良好に表すことができるので音質が向上する。

【0035】

【実施例】次に、本発明について図面を参照して説明する。

【0036】図1は本発明の第1の実施例を示すブロック図である。同図において、入力端子100から入力する音声信号は、フレーム分割回路110でフレーム（例えば40ms）毎に分割され、サブフレーム分割回路120でさらにフレームよりも短いサブフレーム（例えば8ms）に分割される。

【0037】スペクトルパラメータ計算回路200では、少なくとも一つのサブフレームの音声信号に対して、サブフレーム長よりも長い窓（例えば24ms）をかけて音声を取り出してスペクトラパラメータをあらかじめ定められた次数（例えば次数 $P=10$ ）にて計算する。スペクトルパラメータは、特に子音、母音間での過渡区間では時間的に大きく変化するので、短い時間毎に

分析する方が望ましいが、そのようにすると分析に必要な演算量が増大するため、フレーム中のいずれかL個(L)1)のサブフレーム(例えばL=3とし、第1, 3, 5サブフレーム)に対してスペクトルパラメータを計算することにする。そして、分析しなかったサブフレーム(ここでは第2, 4サブフレーム)では、それぞれ、第1と第3サブフレーム、第3と第5サブフレームのスペクトルパラメータを後述のLSP上で直線補間したものをスペクトルパラメータとして使用する。ここでスペクトルパラメータの計算には、周知のLPC分析や、バーク(Burg)分析等を用いることができる。本実施例では、Burg分析を用いる。Burg分析の詳細については、例えば中溝による“信号解析とシステム同定”と題した単行本(コロナ社、1988年刊)の82ないし87ページ(文献6)に記載されている。【0038】さらにスペクトルパラメータ計算回路200では、Burg法により計算された線形予測係数 α_i ($i=1$ ないし10)を量子化や補間に適した線スペクトル対(LSP)パラメータに変換する。ここで、線形予測係数からLSPへの変換法は、例えば菅村らによる“線スペクトル対(LSP)音声分析合成方式による音声情報圧縮”と題した論文(電子通信学会論文誌、J64-A、599ないし606ページ、1981年)(文献7)を利用する。つまり、第1, 3, 5サブフレームでBurg法により求めた線形予測係数を、LSPパラメータに変換し、第2, 4サブフレームのLSPを直線補間により求めて、第2, 4サブフレームのLSPを逆変換して線形予測係数に戻し、第1ないし5サブフレーム

* μ の線形予測係数 α_{i1} ($i=1$ ないし10, $l=1$ ないし5)を聴感重み付け回路230に出力する。また、第1ないし第5サブフレームのLSPをスペクトルパラメータ量子化回路210へ出力する。

【0039】スペクトルパラメータ量子化回路210では、あらかじめ定められたサブフレームのLSPパラメータを効率的に量子化する。本実施例では、量子化法として、ベクトル量子化を用い、第5サブフレームのLSPパラメータを量子化する。LSPパラメータのベクトル量子化の手法は周知の手法を用いることができる。(例えば特願平2-297600号(文献8)や特願平3-261925号(文献9)、特願平3-155049号(文献10)等を参照)さらに、スペクトルパラメータ量子化部210では、第5サブフレームで量子化したLSPパラメータをもとに、第1ないし第4サブフレームのLSPパラメータを復元する。本実施例では、現フレームの第5サブフレームの量子化LSPパラメータと1つ過去のフレームの第5サブフレームの量子化LSPを直線補間して、第1ないし第4サブフレームのLSPを復元する。つまり、量子化前のLSPと量子化後のLSPとの誤差電力を最小化するコードベクトルを1種類選択した後に、直線補間により第1ないし第4サブフレームのLSPを復元できる。さらに性能を向上させるためには、前記誤差電力を最小化するコードベクトルを複数候補選択したのちに、各々の候補について、下式(10)による累積歪を評価し、累積歪を最小化する候補と補間LSPの組を選択するようにすれば良い。

【0040】

$$D = \sum_{l=1}^5 \sum_{i=1}^{10} c_i b_{li} [lsp_{li} - lsp'_{li}]^2 \quad \dots(10)$$

【0041】ここで、 lsp_{li} 、 lsp'_{li} はそれぞれ、第1サブフレームの量子化前のLSP、量子化後復元した第1サブフレームのLSPである。また、 b

* α_{i1} は、第1サブフレームの量子化前のLSPに対して下式(11)を適用して求めた重み係数である。

【0042】

$$b_{li} = (1 / [lsp_{li} - lsp_{li-1}]) + (1 / [lsp_{li+1} - lsp_{li}]) \quad (11)$$

また、 c_i はLSPの次数方向への重み係数であり、例えば下式(12)を用いて求めることができる。★

★【0043】

$$c_i = 1.0 \quad (i=1 \text{ ないし } 8), 0.8 \quad (i=9 \text{ ないし } 10) \quad (12)$$

以上により復元した第1ないし第4サブフレームのLSPと第5サブフレームの量子化LSPをサブフレーム毎に線形予測係数 α'_{i1} ($i=1$ ないし10, $l=1$ ないし5)に変換し、インパルス応答計算回路310へ出力する。また、第5サブフレームの量子化LSPのコードベクトルを表すインデックスをマルチプレクサ400に出力する。

【0044】上記の動作において、直線補間のかわりに、LSPの保管パターンをあらかじめ定められたビット数(例えば2ビット)分用意しておき、これらのパターンの各々に対して第1ないし第4サブフレームのLSP

40 Pを復元して式(10)を評価し、式(10)を最小化するコードベクトルと補間パターンの組を選択するようにしても良い。このようにすると保管パターンのビット数だけ伝送情報が増加するが、LSPのフレーム内での時間的な変化をより精密に表すことができる。ここで、保管パターンは、トレーニング用のLSPデータを用いてあらかじめ学習して作成しても良いし、あらかじめ定められたパターンを格納しておいても良い。

【0045】モード分類回路245では、モード分類を行うための特徴量として、スペクトルパラメータの予測誤差電力を用いる。スペクトルパラメータ計算回路20

11

12

0により計算された線形予測係数を5サブフレーム分入力し、Kパラメータに変換し、下式(13)により5サ

* ブフレーム分の累積予測誤差電力Eを計算する。
【0046】

$$E = 1/5 \sum_{i=1}^5 G_i \quad \dots(13)$$

【0047】ただし、

※ ※ 【0048】

$$G_i = P_i \cdot \left(\prod_{j=1}^{10} [1 - k_{ij}^2] \right) \quad \dots(14)$$

【0049】であり、 P_i は第1サブフレームの入力信号のパワである。次にEの値をあらかじめ定められたしきい値と比較して、複数種類のモードに分類する。例えば4種類のモードに分類するときは、3種類のしきい値と比較することにより行う。分類して得られたモード情報は適応コードブック回路300へ出力されると共に、モード情報を表すインデクス(4種類のモード情報のときは2ビット)がマルチプレクサ400へ出力される。

【0050】重み付け回路230は、スペクトルパラメータ計算回路200から、各サブフレーム毎に線形予測係数 α_{i1} ($i=1$ ないし10、 $1=1$ ないし5)を入力し、式(1)にもとづき、サブフレームの音声信号に対

10★して聴感重み付けを行い、聴感重み付け信号を出力する。

【0051】応答信号計算回路240は、スペクトルパラメータ計算回路200から、各サブフレーム毎に線形予測係数 α_{i1} を入力し、スペクトルパラメータ量子化回路210から、量子化、補間して復元した線形予測係数 α'_{i1} をサブフレーム毎に入力し、保存されているフィルタメモリの値を用いて、入力信号 $d(n)=0$ とした応答信号を1サブフレーム分計算し、減算器250へ出力する。ここで、応答信号 $x_z(n)$ は下式(15)で表される。

【0052】

$$x_z(n) = d(n) - \sum_{i=1}^{10} \alpha_i \cdot d(n-i) + \sum_{i=1}^{10} \alpha_i \gamma^i \cdot y(n-i) + \sum_{i=1}^{10} \alpha'_i \gamma^i \cdot x_z(n-i) \quad \dots(15)$$

【0053】ここで、 γ は式(1)の場合と同一の値である。

【0054】減算器250は、下式により、聴感重み付け信号から応答信号を1サブフレーム分減算し、 $x'_z(n)$ を適応コードブック回路300へ出力する。

【0055】

$$\star x'_z(n) = x_z(n) - x_z(n) \quad (16)$$

インパルス応答計算回路310は、z変換が下式で表される重み付けフィルタのインパルス応答 $h_z(n)$ をあらかじめ定められた点数Lだけ計算し、適応コードブック回路300、音源量子化回路350へ出力する。

☆ 【0056】

$$H_w(z) = [1 - \sum_{i=1}^{10} \alpha_i z^{-i} / 1 - \sum_{i=1}^{10} \alpha_i \gamma^i z^{-i}] \cdot [1 / 1 - \sum_{i=1}^{10} \alpha'_i \gamma^i z^{-i}] \quad \dots(17)$$

【0057】適応コードブック回路300は、モード分類回路からのモード情報を入力し、あらかじめ定められたモードのときのみピッチパラメータを求める。ここでは、モードは4モードあり、モード分類のときのしきい値がモード0からモード3になるにつれて大きくなっているとすると、モード0は子音部、モード1ないし3は母音部に対応すると考えられるので、適応コードブック回路300は、モード1ないし3に対してのみピッチパラメータを求めることにする。まず、オープンループ探索において、聴感重み付け回路230の出力信号に対して、サブフレーム毎に式(2)を最大化する整数遅延候補を複数種類(例えばM種類)選択する。さらに、遅延の短い領域では(例えば遅延が20ないし80)、候補の各々に対して前記文献4等の手法を用いて、整数遅

延の近傍において、複数種類の小数遅延候補を求め、最終的に前記式(2)を最大化する小数遅延候補をサブフレーム毎に少なくとも1種類選択する。以下では、簡単のために候補数は1種類とし、サブフレーム毎に1種類選択された遅延を d_1 ($1=1$ ないし5)とする。次に、クロズドループ探索において、過去のフレームの駆動音源信号 $v(n)$ をもとに、サブフレーム毎に、 d_1 のあらかじめ定められた近傍の点数 ϵ に対して下式(18)を評価し、その値を最大化する遅延をサブフレーム毎に求め、遅延を表すインデクス l_1 をマルチプレクサに出力する。探索法の詳細については、文献5等を参照できる。また、下式(21)により適応コードベクトルを計算し、音源量子化回路350に出力する。

【0058】

$$D' = \sum_{n=0}^{13} (d_1 + \varepsilon) = P' \cdot \sum_{n=0}^{14} (d_1 + \varepsilon) / Q(d_1 + \varepsilon) \quad (18)$$

* * [0059]

ただし

$$P'(d_1 + \varepsilon) = \sum_{n=0}^{N-1} x'_w(n) [v(n - (d_1 + \varepsilon)) * h_w(n)] \quad \dots(19)$$

$$Q(d_1 + \varepsilon) = \sum_{n=0}^{N-1} [v(n - (d_1 + \varepsilon)) * h_w(n)]^2 \quad \dots(20)$$

【0060】ここで、 $h_w(n)$ はインパルス応答計算 10※ [0061] 回路の出力である。

$$q(n) = \beta \cdot v(n - (d_1 + \varepsilon)) * h_w(n) \quad (21)$$

$$\text{ただし、} \beta = P'(d_1 + \varepsilon) / Q(d_1 + \varepsilon) \quad (22)$$

また、作用の項で説明したように、有声区間（例えばモード1ないし3）では、サブフレーム間で遅延の差分をとり、差分を伝送することもできる。このような構成の場合、例えばフレームの第1サブフレームでは、小数遅延で8ビット伝送し、第2～5サブフレームでは前のサブフレームとの遅延の差分を例えばサブフレーム当たり3ビットで伝送することができる。またオープンループ遅延探索のときに、第2～5サブフレームでは、前のフレームの遅延の近傍値を3ビット分探索することとし、さらに各サブフレーム毎に遅延の候補を選択するのではなく、遅延候補の5サブフレーム分のパスに対して5サブフレーム分累積の誤差電力を求め、これを最小化する遅延候補のパスを求め、クローズドループ探索に出力する。クローズドループ探索では、前のサブフレームでク★

★ローズドループ探索により得られた遅延値の近傍を3ビット分探索して最終的な遅延値を求め、求めたサブフレーム毎の遅延値に対応するインデックスをマルチプレクサ400に出力する。

【0062】音源量子化回路350は、減算器250の出力信号、適応コードブック回路300の出力信号、インパルス応答計算回路310の出力信号を入力し、まず、複数段からなるベクトル量子化コードブックの探索を行う。図1では複数種類のベクトル量子化コードブックを音源コードブック351₁～351_mとして表している。以下では簡単のために、段数は2とする。各段のコードベクトルの探索は、式(5)を修正した下式(23)に従う。

【0063】

$$D = \sum_{n=0}^{N-1} [x'_w(n) - \beta \cdot q(n) - \gamma_1 c_{1j}(n) * h_w(n) - \gamma_2 c_{2j}(n) * h_w(n)]^2 \quad \dots(23)$$

【0064】ただし、 $x'_w(n)$ は減算器250の出力信号である。なお、モード0では適応コードブックを使用しないため、式(23)の代わりに、下式(24)★

☆を最小化するコードベクトルを探索する。

【0065】

$$D = \sum_{n=0}^{N-1} [x'_w(n) - \gamma_1 c_{1j}(n) * h_w(n) - \gamma_2 c_{2j}(n) * h_w(n)]^2 \quad \dots(24)$$

【0066】式(23)を最小化するための1段目、2段目のコードベクトルの探索法は種々あるが、ここでは1段目、2段目から複数種類の候補を選択し、その後、両候補の組み合わせ探索を行い、式(23)の値を最小化する候補の組み合わせを決定する。また、1段目、2段目のベクトル量子化コードブックは、前述の探索法を考慮して、あらかじめ多量の音声データベースを用いて設計しておく。以上により決定された1段目、2段目のコードベクトルのインデックス i_{c1} 、 i_{c2} を出力す◆

◆る。

【0067】また、音源量子化回路350は、ゲインコードブックの探索も行う。図1ではゲインコードブックは355として表されている。ゲインコードブックは、適応コードブックを使用するモード1ないし3では、音源コードブックの決定されたインデックスを用いて下式(25)を最小化するように探索する。

【0068】

$$D_k = \sum_{n=0}^{N-1} [x'_w(n) - \beta'_k \cdot q(n) - \gamma'_1 c_{1j}(n) * h_w(n) - \gamma'_2 c_{2j}(n) * h_w(n)]^2 \quad \dots(25)$$

【0069】ここでは、適応コードベクトルのゲイン、1段目、2段目の音源コードベクトルのゲインを3次元のゲインコードブックを使用して量子化することとする。ここで、 $(\beta_k, \gamma_{1k}, \gamma_{2k})$ はそのk番目のコードベクトルである。式(25)を最小化するには、例えば、全てのゲインコードベクトル($k=0$ ないし 2^s-1)に対して式(25)を最小化するゲインコードベクトルを求めてもよいし、ゲインコードベクトルの候補を*

$$D_k = \sum_{n=0}^{N-1} [x'_w(n) - \gamma'_{1k} c_{11}(n) * h_w(n) - \gamma'_{2k} c_{22}(n) * h_w(n)]^2 \quad \cdots (26)$$

【0071】重み付け信号計算回路360は、スペクトルパラメータ計算回路の出力パラメータ及び、それぞれのインデックスを入力し、インデックスからそれに対応する※

$$v(n) = \beta' v(n-d) + \gamma'_1 c_1(n) + \gamma'_2 c_2(n) \quad (27)$$

ただし、適応コードブックを使用しないモードでは、 $\beta' = 0$ とする。次に、スペクトルパラメータ計算回路200の出力パラメータ、スペクトルパラメータ量子化回路210の出力パラメータを用いて下式(28)によし★

$$S_w(n) = v(n) - \sum_{i=0}^{10} a_i v(n-i) + \sum_{i=0}^{10} a_i \gamma'_1 p(n-i) + \sum_{i=0}^{10} a_i \gamma'_2 s_w(n-i) \quad \cdots (28)$$

【0074】以上により、本発明の第1の実施例の説明を終える。

【0075】図2は本発明の第2の実施例を示すブロック図である。本実施例は、モード分類回路410に関するものであり、図1の実施例と同一の参照符号を付した構成要素は、図1の場合と同じ動作を行うので説明を省略する。

* 複数種類予備選択しておき、その複数種類のなかから、式(25)を最小化するものを選択しても良い。ゲインコードベクトル決定後、選択されたゲインコードベクトルを示すインデックス l を出力する。一方、適応コードブックを用いないモードでは、下式(26)を最小化するようにゲインコードブックを探索する。ここでは2次元のゲインコードブックを使用する。

【0070】

※コードベクトルを読みだし、まず下式にもとづき駆動音源信号 $v(n)$ を求める。

【0072】

★り、重み付け信号 $s_w(n)$ をサブフレーム毎に計算し、応答信号計算回路240へ出力する。

【0073】

☆【0076】適応コードブック回路420におけるオープンループ計算回路421は、前述の式(2)および(3)にもとづき、サブフレーム毎に、遅延の候補を少なくとも1種類求め、クローズドループ計算回路422に出力する。さらに、式(29)のピッチ予測誤差電力をサブフレーム1毎に計算する。

☆ 【0077】

$$P_{c1} = \sum_{n=0}^{N-1} x_{w1}^2(n) - P_1^2(T) / Q_1(T) \quad \cdots (29)$$

【0078】そして、 P_{c1} をモード分類回路410に出力する。

【0079】クローズドループ計算回路422は、モード情報、オープンループ計算回路421からサブフレーム毎に少なくとも1種類の遅延候補、聴感重み付け信号を入力し、第1の実施例の適応コードブック回路300◆

◆におけるクローズドループ探索部と同一の動作を行う。

【0080】モード分類回路410は下式(30)に従い、特徴量として、累積ピッチ予測誤差電力 E_p を求め、これを複数種類のしきい値と比較し、モードを分類し、モード情報を出力する。

【0081】

$$E_p = 1 / 5 \sum_{l=1}^5 P_{c1} \quad \cdots (30)$$

【0082】以上で、第2の実施例の説明を終了する。

【0083】図3は本発明の第3の実施例を示すブロック図である。図3において、図1と同一の参照符号を付した構成要素は、図1と同一の動作を行うので、説明は

省略する。図3において、スペクトルパラメータ量子化回路450は、スペクトルパラメータ量子化用に複数種類の量子化コードブック451₁～451_Mを有しており、モード情報分類回路250からモード情報を入力

し、あらかじめ定められたモード毎に451₁、～451_{M-1}を切り替えて使用する。

【0084】量子化コードブック451₁、～451_Mは、あらかじめ、トレーニング用の多量のスペクトルパラメータをモードに分類しておき、あらかじめ定められたモード毎に量子化コードブックを設計しておけばよい。このような構成をとることにより、量子化スペクトルパラメータのインデックスの伝送情報量、コードブック探索の演算量を図1と同一に保ちながら、コードブックサイズが数倍になったこととほぼ等価となるので、スペ

クトルパラメータ量子化の性能を大幅に改善できる。

【0085】以上により、第3の実施例の説明を終了する。

【0086】図4は本発明の第4の実施例を示すブロック図である。。図4において、図1と同一の参照符号を付した構成要素は、図1と同一の動作を行うので、説明は省略する。図4において、音源量子化回路470は、N段(N)1)のベクトル量子化コードブック471₁、～471_MをMセット(M)1)、及びゲインコードブック481を481₁、～481_{M-1}までMセット有しており、471₁、～471_{M-1}(合計N×M種類)のベクトル量子化コードブックを有している。モード分類回路250からのモード情報を用いて、あらかじめ定められたモードのときに、Mセットのうちのあらかじめ定められたj番目のセットにおけるN段のベクトル量子化コードブックを選択し、あらかじめ定められたj番目のセットにおけるゲインコードブックを選択して音源信号の量子化を行う。

【0087】音源コードブック、ゲインコードブックを設計するときは、あらかじめ多量の音声データベースをモード毎に分類しておき、前述の方法を用いて、あらかじめ定められたモード毎にコードブックを設計すれば良い。これらにより、音源コードブック、ゲインコードブックのインデックスの伝送情報量、音現コードブック探索の演算量を図1と同一に保ちながら、コードブックサイズがM倍になったこととほぼ等価となるので、音源量子化の性能を大幅に改善できる。

【0088】図4の音源量子化回路350において、N段のコードブック351₁、～351_Mを有し、そのうちの少なくとも1段が、図5に示すような、あらかじめ定められた間引き率のレギュラーパルス構成となっている。図5では、間引き率m=2の例を示す。レギュラーパルス構成にすると、振巾が零の位置では演算処理が不要なのでコードブック探索に必要な演算量を約1/mに低減化することができる。また、コードブックを格納するのに必要なメモリ量も振巾が零の位置では格納する必

要がないので、約1/mに低減化できる。レギュラーパルス構成の詳細については、デルプラト(M. Delprat)らの“A・6 kbps. レギュラー・パルス・シーイーエルビー・コーダ・フォア・モバイル・ラジオ・コミュニケーションズ(A 6 kbps regular pulse CELP coder for mobile radio communications)”(アタル(Atal)編、Kluwer Academic Publishers, 1990年、179ないし188ページ)と題した論文(文献11)等を参照できるので、ここでは説明を省略する。レギュラーパルス構成のコードブックも前述の方法により、あらかじめ学習しておく。

【0089】さらに、異なる位相の振幅パターンを共通のパターンとして表現してコードブックを設計し、符号化時は、位相だけ時間的にずらして使用することにより、m=2のときは、メモリ量、演算量をさらに1/2に低減化できる。

【0090】また、メモリ量を低減させるためには、レギュラーパルス構成以外にも、マルチパルスの構成をとることもできる。

【0091】以上により本発明の第4の実施例の説明を終了する。

【0092】本発明の意向を損なうこと無く、上述した実施例以外にも種々の変形が可能である。

【0093】まず、スペクトルパラメータはLSP以外にも他の周知なパラメータを用いることができる。

【0094】スペクトルパラメータ計算回路200では、フレーム中で少なくとも1つのサブフレームでスペクトルパラメータを計算するときに、前のサブフレームと現在のサブフレームとのRMSの変化あるいはパワの変化を測定し、これらの変化が大きな複数のサブフレームに対してスペクトルパラメータを計算するようにしてもよい。このようにすると、音声の変化点では必ずスペクトルパラメータを分析することになり、分析するサブフレーム数を低減しても性能の劣化を防ぐことができる。

【0095】スペクトルパラメータの量子化には、ベクトル量子化、スカラー量子化、ベクトル・スカラー量子化など周知な方法を用いることができる。

【0096】スペクトルパラメータ量子化回路における補間パターンの選択には、式(10)以外にも他の周知な距離尺度を用いることができる。例えば、下式(31)を用いることもできる。

【0097】

$$D = \sum_{i=1}^5 R_i \sum_{j=0}^{10} c_{ij} [lsp_{ij} - lsp'_{ij}]^2 \quad \dots(31)$$

$$R_i = RMS_i / \left[\sum_{j=1}^5 RMS_j \right] \quad \dots(32)$$

【0098】ここで、 RMS_i は1サブフレームのRMSもしくはパワーである。

【0099】また、音源量子化回路において、式(23)～(26)で、ゲイン γ_1 と γ_2 を同一とすることもできる。このときは、ゲインコードブックは適応コードブックを用いるモードでは、2次元ゲインとなり、適応コードブックを用いないモードでは1次元ゲインとなる。また、モード毎に、音源コードブックの段数、あるいは各段の音源コードブックのビット数、ゲインコードブックのビット数を変えることもできる。例えば、モード0は3段で、モード1ないし3は2段とすることもできる。

【0100】また、音源コードブックの構成は、例えば、2段構成のときに、1段目コードベクトルに対応させて2段目のコードブックを設計しておき、1段目で選択されたコードベクトルに応じて2段目で探索するコードブックを切り替えるようにすると、メモリ量は増大するが、性能はさらに改善される。

【0101】また、音源コードブックの探索、ならびに学習のときの距離尺度は、他の周知な尺度を用いることもできる。

【0102】また、ゲインコードブックは、伝送ビット数よりも全体で数倍大きなサイズのコードブックをあらかじめ学習し、あらかじめ定められたモード毎に前記コードブックの一部の領域を使用領域としてアサインしておき、符号化するときは、モードに応じて使用領域を切り替えて使用することもできる。

【0103】また、適応コードブック回路での探索、並びに音源量子化回路での探索には、それぞれ、式(19)～(21)、式(23)～(26)のように、インパルス応答 $h_n(n)$ を用いて畳み込み演算を行っているが、これは、伝達特性が式(6)で表されるような重み付けフィルタを用いてフィルタリング演算により行うこともできる。このようにすると、演算量は増大する *

10* が、性能はさらに向上する。

【0104】

【発明の効果】以上述べたように本発明によれば、音声の特徴量を用いて音声をモードに分類し、モードによりスペクトルパラメータの量子化の方法、適応コードブックの動作、音源量子化の方法を切り替えているので、従来方式と比較してより低いビットレートでも良好な音質が得られる。

【図面の簡単な説明】

【図1】本発明の第1の実施例のブロック図。

【図2】本発明の第2の実施例のブロック図。

【図3】本発明の第3の実施例のブロック図。

【図4】本発明の第4の実施例のブロック図。

【図5】本発明の第4の実施例におけるレギュラーパルスを例示する信号タイミング図である。

【符号の説明】

110 フレーム分割回路

120 サブフレーム分割回路

200 スペクトルパラメータ計算回路

210, 450 スペクトルパラメータ量子化回路

211, 451, ..., 451_{M-1} コードブック

230 重み付け回路

240 応答信号計算回路

245, 410 モード分類回路

250 演算回路

300, 420 適応コードブック回路

310 インパルス応答計算回路

350, 470 音源量子化回路

351, ..., 351_N, 471, ..., 471_{M-1} 音源コードブック

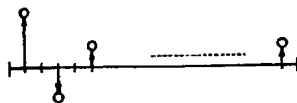
355, 481, ..., 481_{M-1} ゲインコードブック

400 マルチプレクサ

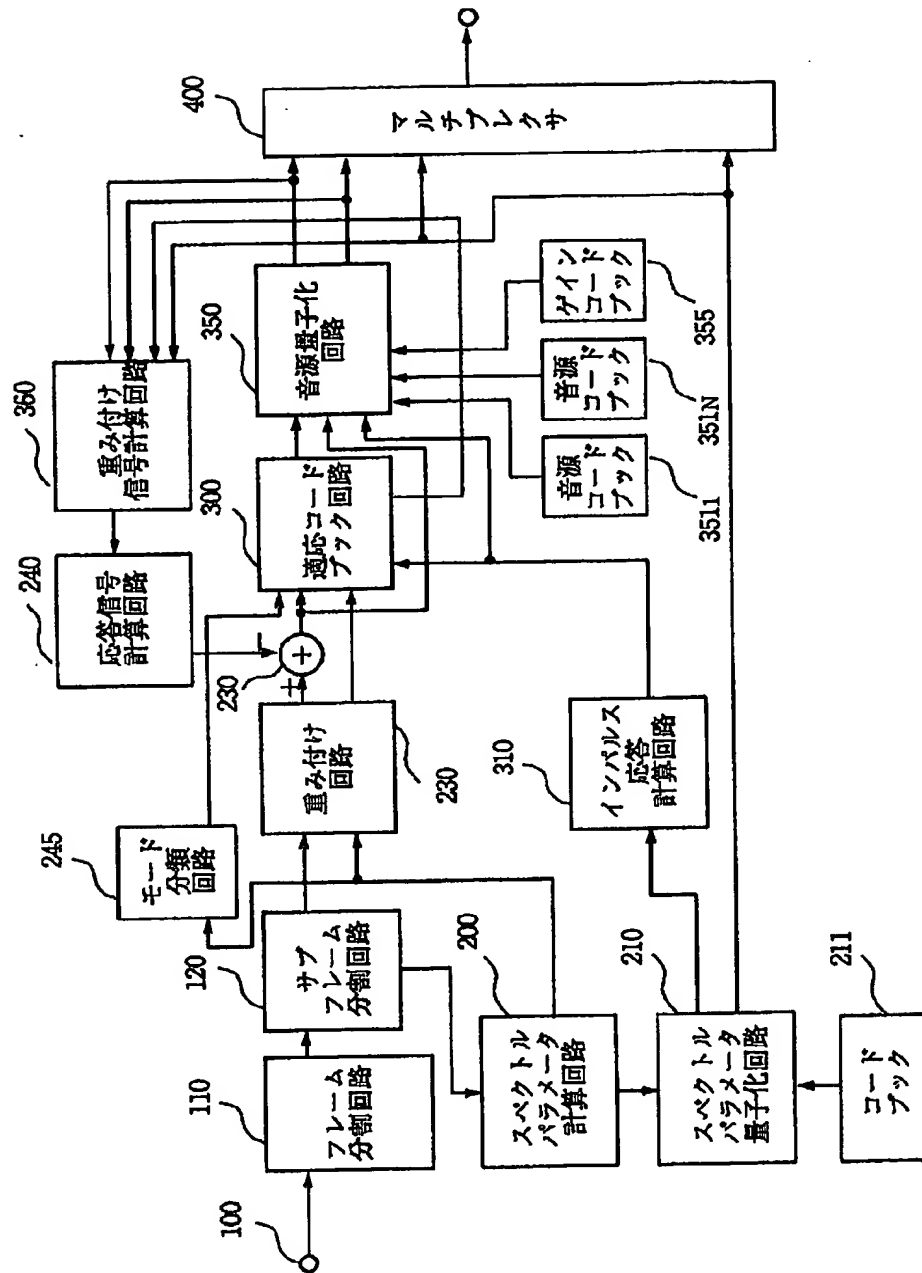
421 オープンループ計算回路

422 クローズドループ計算回路

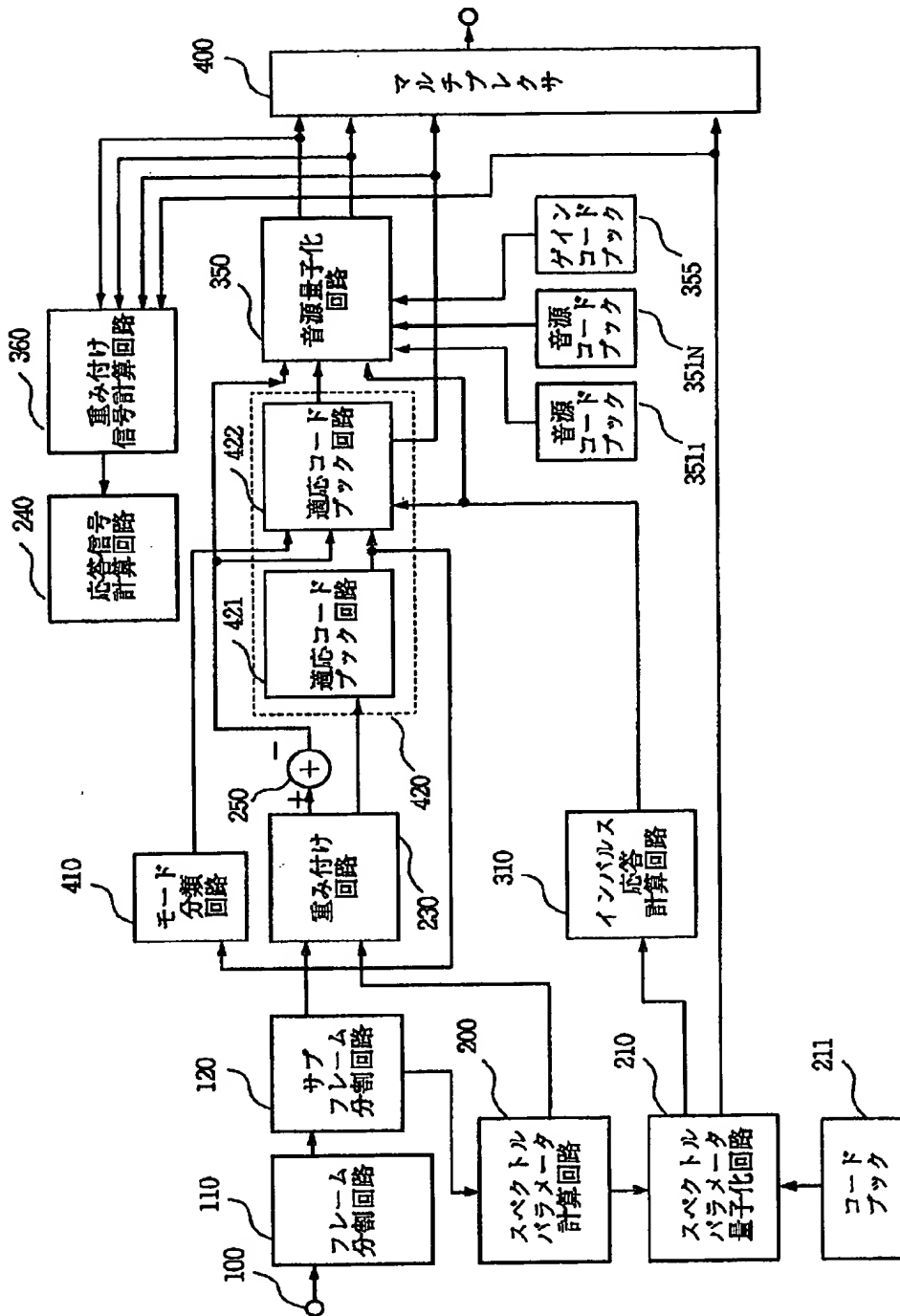
【図5】



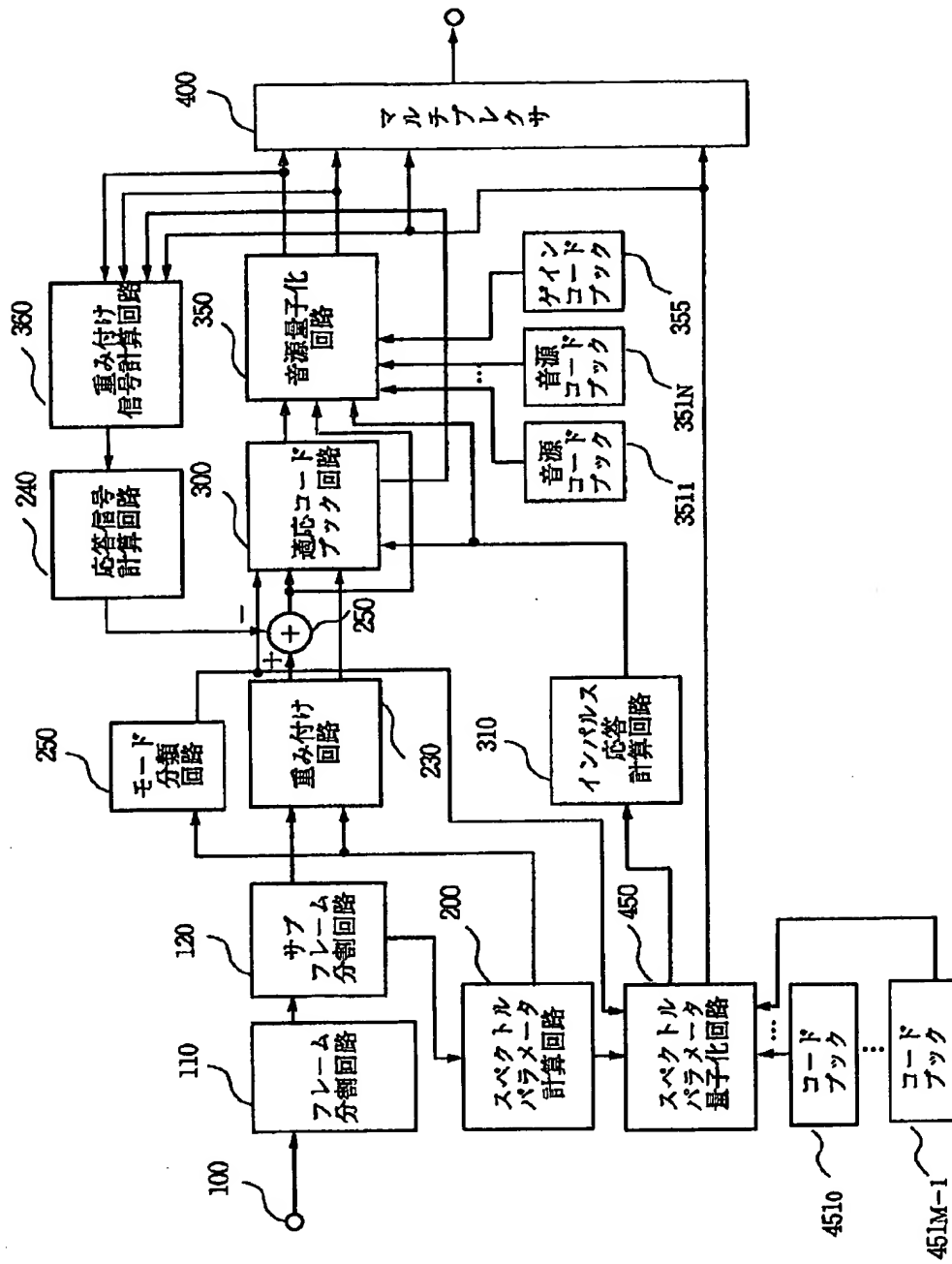
【図1】



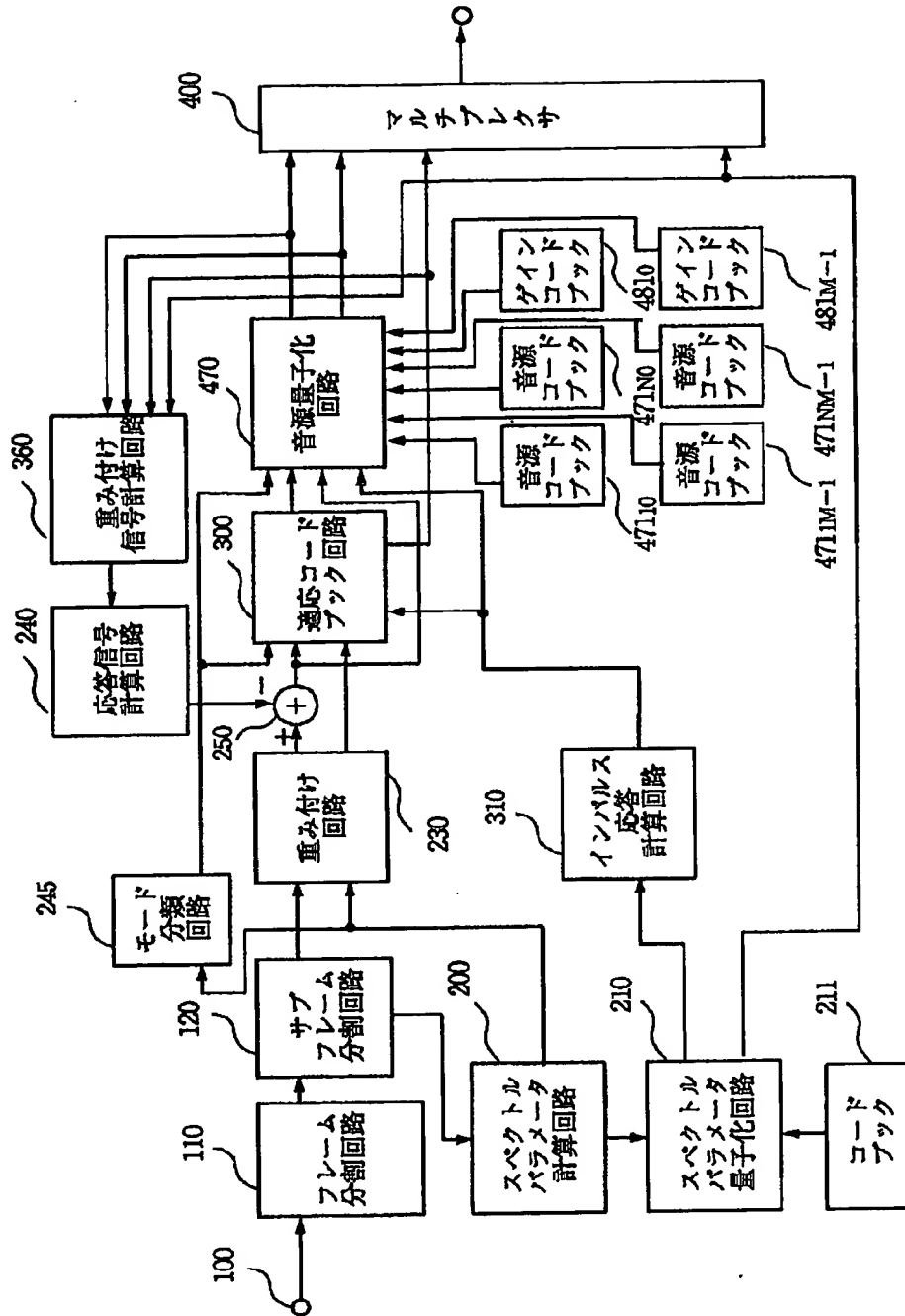
【図2】



【圖3】



Q





PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-222797

(43)Date of publication of application : 12.08.1994

(51)Int.Cl.

G10L 9/14

G10L 9/18

(21)Application number : 05-008737

(71)Applicant : NEC CORP

(22)Date of filing : 22.01.1993

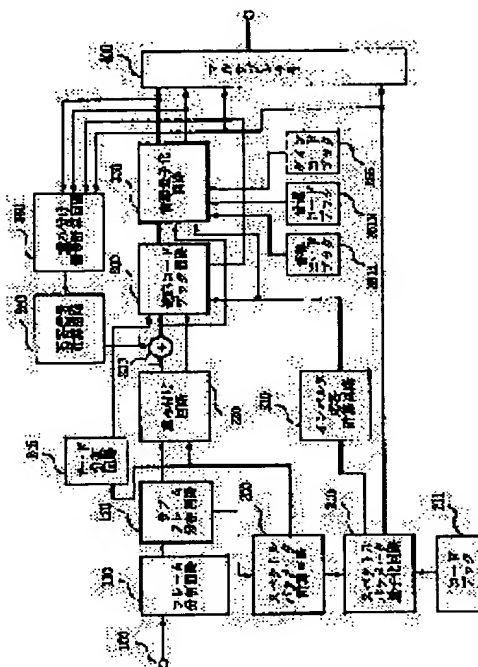
(72)Inventor : OZAWA KAZUNORI

(54) VOICE ENCODING SYSTEM

(57)Abstract:

PURPOSE: To provide an voice encoding system obtaining high tone quality with 4.8kb/s bit rate or below.

CONSTITUTION: This system is provided with a spectrum parameter calculation part 200 dividing an voice signal to frames, dividing it to subframes and calculating the spectrum parameter of the subframe, a spectrum parameter quantizing part 210 quantizing the spectrum parameter, a weighting part 230 calculating a characteristic amount from the voice signal and performing acoustic weighting to the voice signal by using the mode grouping part 245 of the frame voice signal and the spectrum parameter, an adaptive code book part 300 using a mode grouping result and the quantized spectrum parameter and the voice signal, and obtaining a parameter showing a pitch period and a sound source quantizing part 350 using the weighted signal and the output of the adaptive code book part 300 and the spectrum parameter and the quantized spectrum parameter, and retrieving sound source code books 3511-351N consisting of plural stages and a gain code book 355, and quantizing the source signal.



LEGAL STATUS

[Date of request for examination] 22.01.1993

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 2746039

[Date of registration] 13.02.1998

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's]

decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japanese Patent Office